Analysis of a Semiautomatic Algorithm for ECG Heartbeat Classification

M Llamedo[1,2,3], JP Martínez[1,3]

[1]Aragon Inst of Eng Research, IIS Aragón, Univ of Zaragoza, Aragon, Spain
[2]Universidad Tecnológica Nacional, Buenos Aires, Argentina
[3]CIBER of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Spain

Abstract

In this work, we present a semiautomatic algorithm for ECG heartbeat classification, based on a previously developed automatic classifier and a clustering algorithm. Both classifier and clustering algorithms include features from the RR interval series and morphology descriptors calculated from the wavelet transform. Integrating the decisions of both algorithms, the presented algorithm can work automatically or with several degrees of assistance, depending the user expertise. The algorithm was evaluated in the MIT-BIH Arrhythmia database for comparison purposes. In the automatic mode, the algorithm obtained performance figures slightly higher than the original automatic algorithm; but with 5 manually annotated heartbeats in 22 recordings, an improvement of 5% in accuracy (A), global sensitivity (S) and global positive predictive value (P^+) is achieved. For the full-assisted modes the algorithm achieved comparable performance with 55 times less annotation effort, and improved the performance with 42 times less effort. These results represent an improvement in the field of ECG heartbeats classification, concluding that the reference performance can be improved with an efficient use of the assistance provided to the algorithm.

1. Introduction

Cardiovascular diseases are currently the biggest single cause of death in developed countries according to their public health agencies. The analysis of the electrocardiographic signal (ECG) provides a noninvasive and inexpensive technique to analyze the heart function for different cardiac conditions. One important analysis performed in the ECG is the classification of heartbeats, which is important for the study of arrhythmias. The automation of this task is very important for long-term recordings and the detection of subtle arrhythmias.

Many algorithms for ECG heartbeats classification were developed and evaluated in the last decades (see references in [1, 2]) using the available two-lead databases. Some methodological key-points in the development of these classifiers allowed results comparison [1–3]. Probably the most relevant aspects were the use of public databases, the fulfillment of AAMI recommendations [4], the patient-oriented data division [1] and the generalization capability of the classifier [2]. Despite compliance of the enumerated recommendations, the automatic algorithms reviewed continue having issues with performance, generalization or both. Many works addressed this problem and proposed different strategies to improve the performance, as in [5,6]. In the works reviewed, the methodology always involve an expert which provides knowledge to adapt an automatic heartbeat classifier to the ECG under evaluation. The result is an increase in performance at the expense of the automaticity of the classifier.

The objective of this work is to develop and evaluate a semiautomatic algorithm based on previously developed automatic classifier [2], in order to increase its performance with minimum expert assistance. The developed classifier should be useful in both full-automatic or expert-assisted scenarios. The performance will be compared with state of the art algorithms [1, 2, 5].

2. Methods

2.1. ECG database

In this work we used the MIT-BIH Arrhythmia database [7] for training and evaluating the classifier. The database consists of 48 two-lead recordings of approximately 30 minutes and sampled at 360 Hz. The annotations provided with the database were used for training and testing purposes, following the recommendations and class-labeling of AAMI. We adopted the same data division used in [1, 5] for comparative purposes. Also AAMI unclassified class (Q) was discarded since it is poorly represented in the database. Finally, a class-labeling modification to the AAMI standard was evaluated, considering fusion (of normal and ventricular beats) and ventricular classes, as the same extended ventricular class (V'). We will refer to this modification as AAMI2 labeling. The division scheme is the same used in other works, and is summarized in Table 1.

Table 1. Scheme of the division of the MIT-BIH database into training (DS1) and testing (DS2) sets.

	N	S	V	F	#Rec
$DS1$	45673	929	3755	412	22
$DS2$	44053	1833	3202	388	22
Full MIT-BIH	88175	1635	7121	822	44

Heart beats classes are N: normal, S: supraventricular, V: ventricular and F: fusion. $DS1$ comprises recordings 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230.

$DS2$ comprises recordings 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234.

2.2. Heartbeats classification

The semiautomatic algorithm includes a linear discriminant classifier (LDC) and a clustering algorithm based on expectation-maximization (EMC), both perform a preliminary classification. The LDC was developed and trained as described in [2]. Both the heartbeat and cluster labels provided by the LDC and EMC respectively, are integrated into a final heartbeat label. This label integration can be performed in three ways, depending on the degree of expert assistance required in the application scenario. The modes of operation are 1) full-automatic, 2) slightly semiautomatic assisted and 3) assisted semiautomatic. For all the modes, the algorithm performs the following procedures: a) Cluster discovery and centroid identification (by computing EMC), b) LDC automatic classification and c) expert assistance.

For mode 1, procedures a) and b) are executed. The result of a) is the classification of the heartbeats into K clusters. Then for each cluster the LDC classifies the heartbeats included, and if the most represented class exceed half of the cluster population the same label is asigned to all this cluster. In case not exceeding the threshold, the LDC labels remain the same. Mode 2 is similar to 1, with the exception that in case not exceeding the population threshold, expert assistance is required to label the cluster centroid and propagate it to the cluster. The procedure of expert assistance is simulated by inspecting the true labels provided with the database. In mode 3, only procedures a) and c) are executed. As a result, the expert task is to label each centroid example provided by a), the algorithm assigns this label to the rest of examples in each cluster.

The LDC classifier used is used under the assumption of independent and normally distributed data, the maximum *a posteriori* criterion (MAP) leads to the quadratic classifier defined by the discriminant functions

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{i}}^{-1}\mathbf{x} + \boldsymbol{\mu}_{\mathbf{i}}^{\mathbf{T}}\boldsymbol{\Sigma}_{\mathbf{i}}^{-1}\mathbf{x} - \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\mu}_{\mathbf{i}}^{\mathbf{T}}\boldsymbol{\Sigma}_{\mathbf{i}}^{-1}\boldsymbol{\mu}_{\mathbf{i}} \\ &\quad -\frac{1}{2}\log(|\boldsymbol{\Sigma}_i|) + \log(P(\omega_i)) \end{aligned}$$

Table 2. Features used in the LDC model obtained in [2].

Feature	Description
$\ln(RR[i])$	Current RR interval
$\ln(RR[i+1])$	Next RR interval
$\ln(RR_1)$	Average RR interval in the last minute
$\ln(RR_{20})$	Average RR interval in the last 20 minutes
$\ln(k_Z^1)$	Zero-cross position of the WT autocorrelation sequence in lead 1
$\ln(k_Z^2)$	Zero-cross position of the WT autocorrelation sequence in lead 2
k_M^1	Maximum position of the WT autocorrelation sequence in lead 1
k_M^2	Maximum position of the WT autocorrelation sequence in lead 2

for the i-th class, where $\mathbf{x}$ represents the feature vector describing each heartbeat, and $\boldsymbol{\mu}_{\mathbf{i}}$, $\boldsymbol{\Sigma}_i$ and $P(\omega_i)$ are the mean vector, covariance matrix and prior probability of the i-th class. The values of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ were computed from the training data with the sample mean and covariance matrix expressions while the values for the prior probabilities $P(\omega_i)$ were considered the same for all classes. The classification rule assigns $\mathbf{x}$ to the class i which results in the maximum posterior probability $g_i(\mathbf{x})$. In the case that the covariance matrix $\boldsymbol{\Sigma}$ is assumed to be the same for all classes ($\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_j, \forall i \neq j$), the quadratic discriminant classifier (QDC) becomes linear in $\mathbf{x}$ leading to the linear discriminant classifier (LDC) where $\boldsymbol{\Sigma}$ can be estimated as the weighted sample covariance

$$\boldsymbol{\Sigma} = \frac{\sum_{i=1}^{C} w_i \sum_{m=1}^{M_i}(\mathbf{x_m} - \boldsymbol{\mu}_{\mathbf{i}}).(\mathbf{x_m} - \boldsymbol{\mu}_{\mathbf{i}})^{\mathrm{T}}}{\sum_{i=1}^{C} w_i.M_i}.$$

The class-weighting possibility is of much interest due to the heavy class-size unbalance inherent to this application, where the normal class is in general one order of magnitude more represented that other classes. In this work, all classification tasks were performed using and adapting the PRtools toolbox [8] for Matlab (The Mathworks Inc., Massachusetts).

In [2] we developed a classification model with good generalization capabilities including rhythm and morphological features. In this work we use the same classification model, which includes the features described in Table 2.

The EMC algorithm used in this work is based in the mixture of Gaussians model [9]. It consists in estimating the parameters of a density function modeled by

$$p(\mathbf{x}|\boldsymbol{\Psi}) = \sum_{\mathbf{k=1}}^{\mathbf{K}} \pi_{\mathbf{k}} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{k}}, \boldsymbol{\Sigma}_{\mathbf{k}}),$$

where K Gaussians are mixed with the coefficient π_k to retain a more realistic structure of the data. Considering the parameter set $\Psi = \{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k | k = 1, \ldots, K\}$, one method to calculate Ψ is by the maximum likelihood estimation to optimize the log likelihood

$$L(X|\Psi) \quad = \quad \ln \prod_{n=1}^{N} p(\mathbf{x}|\boldsymbol{\Psi}),$$

Table 3. Features used with the EMC algorithm.

Feature	# features	Description
$\ln(RR[i-1])$	1	Previous RR interval
$\ln(P_{RR})$	2	Prematurity of the heartbeat
$\ln(dRR_L)$	1	Local RR interval variation
$\ln(RR_{1,5})$	2	Mean RR interval within the last 1 and 5 minutes
$\ln(\sigma_{RR_{10}})$	1	RR interval standard deviation within the last 10 minutes
$\ln(QRS_W)$	1	QRS width measured with [10]
$\ln(S^{1,2}_{QRS})$	2	QRS mean wavelet scale at leads 1 and 2
$\ln(t^1_2)$	1	Position of the second maximum of 4th scale WT of the QRS complex at lead 1
$W_4x(t^{1,2}_1)$	2	Value of the first maximum of the 4th scale WT of the QRS complex at lead 1 and 2
$r_T(k^1_M)$	1	Value of the first maximum in the T wave autocorr. sequence at lead 1
$r_{QRST}(k^1_M)$	1	Value of the first maximum in the QRST complex autocorr. sequence at lead 1

for the N heartbeats in each recording named $X = \{\mathbf{x_1}, \ldots, \mathbf{x_N}\}$. Since there is not a closed form solution for Ψ by optimizing $L(X|\Psi)$, the well-known expectation-maximization algorithm (EM) is used to obtain the estimation equations of the parameters Ψ [9]. The interested reader is referred to [8, 9] for details, equations and the implementation used in this work.

Regarding to the feature space used with the EMC, we followed the same feature selection procedure described in [2], by means of a sequential floating feature selection algorithm (SFFS). For the case of clustering instead of looking for features with generalization capability or interpatient separability, we looked for those with intrapatient separability. As a result a model of 15 features was obtained, this model also includes a description of the rhythm and morphology of heartbeats as can be seen in Table 3.

Among the rhythm features used in the model, we have two measures of the prematurity of a heartbeat

$$P^1_{RR}[i] = \frac{RR[i]}{\sum_{k=i-1}^{i+1} RR[k]}$$

and

$$P^2_{RR}[i] = \frac{RR[i]}{RR[i] + RR_{30s} + \max_{10s}(RR)}.$$

The first measures the prematurity against the previous and next RR interval, while the second does relative to the maximum RR interval in 10 seconds and to the mean in 30 seconds. The local RR interval variation is defined as $dRR_L[i] = \sum_{k=i-1}^{i+1} |dRR[k]|$, where $dRR[i] = RR[i] - RR[i-1]$. One of the morphology related feature is the wavelet scale where the QRS complex is mostly projected, since fast evolving signals (like a normal beat) tend to be projected in lower wavelet scales (higher frequency content). This feature is calculated as a weighted sum, where the

$$A_i = \frac{1}{M} \sum_{m=1}^{2} |W_i x(t_m)|$$

are the mean absolute peak amplitude for scales $i = 1, 2, \ldots 6$, being 2 the number of peaks detected at each scale at times t_m. Then is calculated the QRS projected scale for each lead (S^{lead}_{QRS}) as

$$S_{QRS} = \frac{\sum_{i=1}^{6} A_i.i}{\sum_{i=1}^{6} A_i}.$$

Other morphologic features are the position of the k-th maximum of the 4th scale of the wavelet transform (WT) at lead l (t^l_k), or the value at this position ($W_4x(t^l_k)$). The last of the morphologic features are the maximum of the autocorrelation sequence calculated at scale 4 of the WT ($r_C(k^l_M)$), similarly to $r(k^l_M)$ calculated in [2]. In this work C is the complex of waves (T, QRS, QRST) and l the lead where the feature is measured.

For the performance evaluation of the algorithm we follow the methodology presented in [2]. As the initialization of the EMC is random, the results of the clustering algorithm are not deterministic. Then each experiment is repeated 30 times to evaluate the center and dispersion of the performance estimates, median and median absolute deviation (MAD) respectively. We are also interested in evaluating the amount of expert assistance required in the semiautomatic modes of operation.

3. Results

In this work we evaluated the semiautomatic algorithm in DS2 of the MITBIH-AR, for the three possible operating modes. The results of this experiments are presented in table 4 and compared with the results obtained in [2, 5].

4. Discussion and conclusions

In this work we presented a versatile ECG heartbeat classification algorithm useful in a broad range of scenarios, from unattended to fully expert-assisted mode. The automatic part of the algorithm relies in a previously developed automatic classifier with proven generalization capability [2]. The assisted part is based on a cluster algorithm, responsible of retaining most of the patient specific features of the heartbeats. For this reason the features model that should be used for the cluster algorithm pursues the maximum intrapatient class separability. This approach is different to the one used in the development of the ELDC feature set in [2]. The complete pool of features used in this work for developing the EMC was the same used in [2], therefore a careful design and selection of other features may improve the performance.

Table 4. Median performances and MAD dispersions in percentages, evaluated 30 times in the 22 recordings of DS2. Comparison between the several operating modes separating AAMI2 classes (N, S, V') with different amount of expert assistance.

				Normal		Suprav.		Ventricular		Total		
Operation mode	Observation	# Clusters	# MAHB/R	S	P^+	S	P^+	S	P^+	A	S	P^+
Semiautomatic	de Chazal 2006	–	500	94	89	88	93	95	96	92	92	93
	FA	12	12±0	99±0	86±2	87±3	99±0	95±1	99±0	94±1	94±1	95±1
		9	9±0	99±0	84±4	86±4	99±0	94±1	99±0	93±1	93±1	94±1
	SA	12	0.3±0	97±0	80±2	81±4	96±0	88±3	93±5	89±2	89±2	90±1
		9	0.3±0	97±1	78±3	82±4	96±1	89±3	99±0	89±2	89±2	91±1
Automatic		12	0	97±0	78±3	77±6	95±1	87±3	92±4	87±2	87±2	89±2
		9	0	97±1	73±3	74±6	95±1	84±4	96±1	85±2	85±2	88±1
	Llamedo 2011	–	0	95	79	77	88	81	88	84	84	85

MAHB/R: manually annotated heartbeats per recording. FA: fully assisted. SA: slightly assisted.

The algorithm has the possibility of graduating the expert assistance from zero to a completely assisted mode. The automatic mode achieved performance figures slightly higher than the obtained in [2]. This is an interesting result since the semiautomatic algorithm without any assistance improves the performance of the LDC alone, from around 84% for A, S and P^+ to more than 85%, as can be seen in Table 4. With a small degree of assistance, 5 manually annotated heartbeats (MAHB) in 22 recordings, these figures increase to more than 89% for the same estimates. This experiment evidences that the algorithm can handle properly the assistance given by an expert. However, the performance achieved might be dependent of the dataset used, as we have previously shown in [2], therefore the evaluation of this algorithm should be extended outside the MITBIH-AR.

In the completely assisted mode, the algorithm needs 9 MAHB per record (MAHB/R) to achieve the same performance than [5], this represents an effort in the manual annotation task 55 times smaller (500 MAHB/R). Besides, annotating 12 beats per record the algorithm can improve the performance reported in [5], but with 42 times less effort.

These results represent an improvement in the field of automatic and semiautomatic heartbeats classification respect to the reference approaches [2, 5], concluding that the performance in [2] can be improved with an efficient handling of the expert assistance.

Acknowledgments

This work was supported by projects TEC2010-21703-C03-02 from MICINN (Spain) and GTC T-30 from DGA (Spain). The CIBER of Bioengineering, Biomaterials and Nanomedicine is an initiative of ISCIII.

References

[1] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2004;51:1196–1206.

[2] Llamedo M, Martínez J. Heartbeat classification using feature selection driven by database generalization criteria. IEEE Transactions on Biomedical Engineering 2011; 58:616–625.

[3] Park K, Cho B, Lee D, Song S, Lee J. Hierarchical support vector machine. In Computers in Cardiology 2008, volume 35. IEEE Computer Society Press, 2008; 229–232.

[4] Testing and reporting performance results of cardiac rhythm and st-segment measurement algorithms. American National Standard, ANSI/AAMI/ISO EC57, 1998–(R)2008.

[5] de Chazal P, Reilly RB. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2006; 53:2535–2543.

[6] Hu YH, Palreddy S, Tompkins W. A patient-adaptable ecg beat classifier using mixture of experts approach. IEEE Transactions on Biomedical Engineering 1997;44:891–899.

[7] Moody G, Mark R. The impact of the MIT-BIH arrhythmia database. IEEE Eng in Med and Biol May-June 2001; 20(3):45–50.

[8] Duin R, Juszczak P, Paclik P, Pekalska E, deRidder D, Tax D, Verzakov S. Pr-tools, a matlab toolbox for pattern recognition, 2008. URL `http://www.prtools.org`.

[9] van der Heijden F, Duin R, de Ridder D, Tax D. Classification, Parameter Estimation and State Estimation. John Wiley & Sons, 2005.

[10] Martínez JP, Almeida R, Olmos S, Rocha A, Laguna P. A wavelet-based ecg delineator: Evaluation on standard databases. IEEE Transactions on Biomedical Engineering 2004;51:570–581.

Address for correspondence:

Mariano Llamedo Soria, llamedom@unizar.es
I3A - I+D Building, C/ Mariano Esquillor S/N
Despacho 4.0.05 – CP: 50018, Zaragoza, España.